# Search Strategies for IR Spectra - Part III
# Increasing the Information of the Search Results

Author: Michael Boruta

This Application Note will address common questions that we receive regarding IR database searching. Our purpose is to help you with your selection of algorithms and the interpretation of your results. This note will explain useful techniques which can be employed to increase the information of the search results. The features and techniques described in this note are available in Bio-Rad's search software. Other software packages which utilize Bio-Rad spectral databases may or may not fully support these features.

Bio-Rad's commercial compound databases contain information about the physical properties of the compounds. The databases generally contain only the information supplied to us by the manufacturer of the product. The amount of information provided in these databases varies from supplier to supplier and sometimes from product to product.

The chemical descriptions are provided by the manufacturer and are not the results of any analysis performed at Bio-Rad. Compounds in these databases are usually named by their trade names. Any chemical names or descriptions are usually found within the classification or comment fields of the physical properties option.

What this means is that if you are searching for quaternary ammonium salts and you only perform a name search, you are likely to miss many if not all of these compounds. For example, if you were to perform a name search for “quaternary ammonium” using the Surfactants Select IR Database, Volume 2, you would not get any matches. Yet, if you performed a physical property search for “quaternary ammonium” in the classification or the comment field, you would find 94 compounds which matched. Most of these matches would also contain additional information as to their chemical composition.

Sometimes when viewing the physical property information available for the compounds in your hit list, you will not get as much information back as you would like. Because we are aware of the limitations of the information supplied by the manufacturers, some techniques have been developed which often provide additional information.

The simplest technique is to look at the physical property information for several similar compounds in the hit list, instead of reviewing just the first hit. Another technique used when facing a reference spectrum which does not provide sufficient information, or when trying to verify the information provided by the manufacturer, is to search a database spectrum itself against the database. For example, in the Surfactants Select IR Database, Volume 2, the compound CM502 (AGEFLEX FM-1Q-80) is described as a quaternary ammonium salt. Searching the database spectrum of CM502 against the same database it came from gave a hit list in which 8 out of the top 10 hits were described as sulfated and sulfonated carboxylic acid esters.

Another approach which one might use is to search the database spectrum against another database which contains more chemical information or structures. This approach can be helpful in classifying bands in a compound, even when a good match from the database is not expected.

Searching a database entry against another database which has structural information can provide useful information, even when the two databases are not closely related.

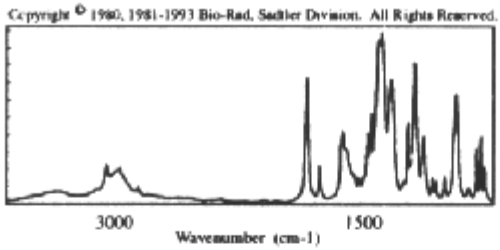
Figure 1 shows the results of using the database entry from the previous example, CM502, and searching its spectrum against the Basic Monomers and Polymers Select IR Database Volume 2, which has structures. Surprisingly, the first hit, BM807, appears to be an exact match with the “unknown” CM502. These two spectra are of the same material from different sources.

Finding an exact match would obviously be considered an exception rather than the rule. Most chemists would agree that the second hit would also be very useful in classifying the compound.

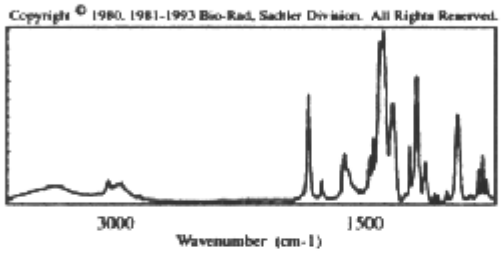
To summarize, when using Bio-Rad's commercial databases, it is best to use the classification and the comment fields when searching for chemical information. Using the name search will not work well because most of the data is listed by trade name. In addition, using a closely matching database entry as the search query can also be a useful tool for obtaining additional chemical or structural information.

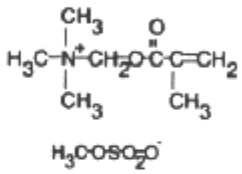
*Figure 1.*

**CM502; AGEFLEX FM-1Q-80**

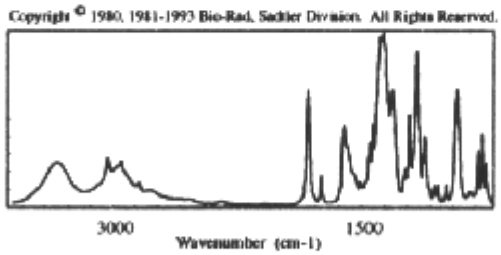


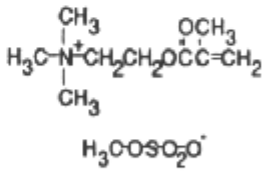

**BM807; TRIMETHYLAMMONIUMMETHYL METHACRYLATE METHOSULFATE, 40% AQUEOUS**



**BM301; METHACRYLYL CHOLINE METHYL SULFATE, 40% AQUEOUS**



D:\TESTDATA\GEFTY\CM502.HIT

| | |
|---|---|
| SEARCH TYPE: | IR Spectrum Search |
| SOURCE: | CM502 |
| DATE: | 7/12/94 |
| TIME: | 22:05:57 |
| NUMBER OF HITS: | 10 |
| SEARCH ALGORITHM: | Euclidean Distance |

| RANK | LIBENTRY | HQI | CHEMICAL NAME |
|---|---|---|---|
| 1 | BM807 | 975 | TRIMETHYLAMMONIUMMETHYL METHACRYLATE METHOSULFATE, 40% AQUEOUS |
| 2 | BM301 | 974 | METHACRYLYL CHOLINE METHYL SULFATE, 40 % AQUEOUS |
| 3 | BM564 | 865 | 2-PHENOXYETHYL METHACRYLATE |
| 4 | BM496 | 841 | POLY(ETHYL) ACRYLATE) |


BIO-RAD *Bio-Rad Laboratories*

***Informatics Division***
*www.sadtler.com*

***Sadtler Software & Databases***

***U.S. Sales*** *Phone: 888-5-BIO-RAD (888-524-6723) • E-mail: sadtler.usa@bio-rad.com*
***Europe*** *Phone: +44 20 8328 2555 • Free phone: 00800 78945000 • E-mail: sadtler.europe@bio-rad.com*
***Japan*** *Phone: +81 03 (5811) 6287 • E-mail: sadtler_nbr@jp.bio-rad.com*
***Rest of World*** *Phone: 215-382-7800 • E-mail: sadtler.row@bio-rad.com*